

# Welcome to the first MMTx User's Group Meeting

*AMIA 2003*

*November 11, 2003*



U.S. National Library of Medicine

# AGENDA

- History
- Latest Release (V2.3.A)
- Customizing Data (Data File Builder)
- Customizing Behavior (MMTx API)
- Future Direction
- Open Session
- Our Questions to You
- Relevant Links



U.S. National Library of Medicine

# History: What is MMTx?

**MetaMap Transfer (MMTx)** is an effort to make the MetaMap program available to biomedical researchers in a generic, configurable environment.

MetaMap maps arbitrary text to concepts in the UMLS Metathesaurus; or, equivalently, it discovers Metathesaurus concepts in text.



U. S. National Library of Medicine

# History: How does MetaMap/MMTx Work?

- Text is processed through a series of modules
- First it is parsed into components including sentences, paragraphs, phrases, lexical elements and tokens.
- Variants are generated from the resulting phrases.
- Candidate concepts from the UMLS Metathesaurus are retrieved and evaluated against the phrases.
- The best of the candidates are then organized into a final mapping in such a way as to best cover the text.



U.S. National Library of Medicine

# History: MMTx Development Team

Alan Aronson (Project Leader & Developer of MetaMap)

## Team Members:

- Guy Divita
- Jim Mork
- Cliff Gay
- Willie Rogers



U.S. National Library of Medicine

# History: MMTx Evolution

- V1.0.A July 23, 2001 (Beta Release)
- V2.0.A February 15, 2002 (Data File Builder & Source)
- V2.0.C March 18, 2002 (Trouble Reporting & Tracking)
- V2.2.A September 16, 2002 (MMTx API)
- V2.3.A October 20, 2003



U.S. National Library of Medicine

# Latest Release: V2.3.A

- Released October 20, 2003
- Quick browse mode
- Single jar file to ease incorporating MMTx into your applications
- Metathesaurus tables reorganized to improve performance
- 2003 Data Models



U.S. National Library of Medicine

# Customizing Data: Data File Builder

- Allows replacement of standard data files used by MMTx
  - This changes the range of concepts to which MMTx maps free text.
- Customize the target concepts when ...
  - You have Copyright issues with the full Metathesaurus
  - You need to apply MMTx to a new domain
  - You want to add a local vocabulary



U.S. National Library of Medicine



# Customizing Data: Using Data File Builder

- Three stages:
  - Create custom Metathesaurus.
    - Use MetamorphoSys or create pseudo-metathesaurus files
  - Generate MMTx data files.
    - Data File Builder scripts and programs automate this step.
    - `mmtx/config/rundatafiles.sh` will reduce the manual steps.
  - Load custom data files to new databases.
    - This will enable running MMTx with a new `mm_data_version`



U.S. National Library of Medicine

# Customizing Behavior: MMTx API

- The MMTx API is the recommended way to embed MMTx within other applications.
- Use the API to
  - Integrate MMTx mapping into larger applications
  - Customize MMTx output
  - Get output for each phrase
  - Annotate your own text with concepts
  - Make your own improvements
- To use the API you need to understand the
  - Parts List (Document Model) and the
  - Process Methods



U.S. National Library of Medicine

# MMTx API: Parts List (Document Model)

- Sections
- Sentences
- Phrases
- Terms
- Words

**SPME determination of volatile aldehydes for evaluation of in-vitro antioxidant activity**

Elena E. Stashenko, Miguel A. Puertas, Jairo R. Martínez  
A<sup>1</sup> Chromatography Laboratory, Research Center for Biomolecules,  
School of Sciences, Industrial University of Santander. A.A. 678,  
Bucaramanga, Colombia

**Abstract:**

The in-vitro antioxidant activity of natural (essential oils, vitamin E) or synthetic substances (tert-butyl hydroxy anisole (BHA), Trolox) has been evaluated by monitoring volatile carbonyl compounds released in model lipid systems subjected to peroxidation. The procedure employed methodology previously developed for the determination of carbonyl compounds as their pentafluorophenylhydrazine derivatives which were quantified, with high sensitivity, by means of capillary gas chromatography with electron-capture detection. Linoleic acid and sunflower oil were used as model lipid systems. Lipid peroxidation was induced in linoleic acid by the Fe<sup>2+</sup> ion (1 mmol L<sup>-1</sup>, 37 °C, 12 h) and in sunflower oil by heating in the presence of O<sub>2</sub> (220 °C, 2 h).

Title

Abstract



U.S. National Library of Medicine

# MMTx API: Parts List (Document Model)

- Sections
- Sentences
- Phrases
- Terms
- Words

## SPME determination of volatile aldehydes for evaluation of in-vitro antioxidant activity

Elena E. Stashenko, Miguel A. Puertas, Jairo R. Martínez  
A<sup>1</sup> Chromatography Laboratory, Research Center for Biomolecules,  
School of Sciences, Industrial University of Santander. A.A. 678,  
Bucaramanga, Colombia

### Abstract:

The in-vitro antioxidant activity of natural (essential oils, vitamin E) or synthetic substances (tert-butyl hydroxy anisole (BHA), Trolox) has been evaluated by monitoring volatile carbonyl compounds released in model lipid systems subjected to peroxidation. The procedure employed methodology previously developed for the determination of carbonyl compounds as their pentafluorophenylhydrazine derivatives which were quantified, with high sensitivity, by means of capillary gas chromatography with electron-capture detection. Linoleic acid and sunflower oil were used as model lipid systems. Lipid peroxidation was induced in linoleic acid by the Fe<sup>2+</sup> ion (1 mmol L<sup>-1</sup>, 37 °C, 12 h) and in sunflower oil by heating in the presence of O<sub>2</sub> (220 °C, 2 h).



U.S. National Library of Medicine

# MMTx API: Parts List (Document Model)

- Sections
- Sentences
- Phrases
- Terms
- Words

## SPME determination of volatile aldehydes for evaluation of in-vitro antioxidant activity

Elena E. Stashenko, Miguel A. Puertas, Jairo R. Martínez  
A1 Chromatography Laboratory, Research Center for Biomolecules,  
School of Sciences, Industrial University of Santander. A.A. 678,  
Bucaramanga, Colombia

### Abstract:

The in-vitro antioxidant activity of natural (essential oils, vitamin E) or synthetic substances (tert-butyl hydroxy anisole (BHA), Trolox) has been evaluated by monitoring volatile carbonyl compounds released in model lipid systems subjected to peroxidation. The procedure employed methodology previously developed for the determination of carbonyl compounds as their pentafluorophenylhydrazine derivatives which were quantified, with high sensitivity, by means of capillary gas chromatography with electron-capture detection. Linoleic acid and sunflower oil were used as model lipid systems. Lipid peroxidation was induced in linoleic acid by the  $\text{Fe}^{2+}$  ion ( $1 \text{ mmol L}^{-1}$ ,  $37^\circ\text{C}$ , 12 h) and in sunflower oil by heating in the presence of  $\text{O}_2$  ( $220^\circ\text{C}$ , 2 h).



U.S. National Library of Medicine



# MMTx API: Parts List (Document Model)

- Sections
- Sentences
- Phrases
- Terms
- Words

## SPME determination of volatile aldehydes for evaluation of in-vitro antioxidant activity

Elena E. Stashenko, Miguel A. Puertas, Jairo R. Martínez  
A<sup>1</sup> Chromatography Laboratory, Research Center for Biomolecules,  
School of Sciences, Industrial University of Santander. A.A. 678,  
Bucaramanga, Colombia

### Abstract:

The in-vitro antioxidant activity of natural (essential oils, vitamin E) or synthetic substances (tert-butyl hydroxy anisole (BHA), Trolox) has been evaluated by monitoring volatile carbonyl compounds released in model lipid systems subjected to peroxidation. The procedure employed methodology previously developed for the determination of carbonyl compounds as their pentafluorophenylhydrazine derivatives which were quantified, with high sensitivity, by means of capillary gas chromatography with electron capture detection. Linoleic acid and sunflower oil were used as model lipid systems. Lipid peroxidation was induced in linoleic acid by the Fe<sup>2+</sup> ion (1 nmol L<sup>-1</sup>, 37 °C, 12 h) and in sunflower oil by heating in the presence of O<sub>2</sub> (220 °C, 2 h).

single word term  
Multi-word term



U.S. National Library of Medicine

# MMTx API: Parts List (Document Model)

- Sections
- Sentences
- Phrases
- Terms
- Words →

## SPME determination of volatile aldehydes for evaluation of in-vitro antioxidant activity

Elena E. Stashenko, Miguel A. Puertas, Jairo R. Martínez  
A<sup>1</sup> Chromatography Laboratory, Research Center for Biomolecules,  
School of Sciences, Industrial University of Santander. A.A. 678,  
Bucaramanga, Colombia

### Abstract:

The in vitro antioxidant activity of natural essential oils, vitamin E or synthetic substances (tert-butyl hydroxy anisole (BHA), Trolox) has been evaluated by monitoring volatile carbonyl compounds released in model lipid systems subjected to peroxidation. The procedure employed methodology previously developed for the determination of carbonyl compounds as their pentafluorophenylhydrazine derivatives which were quantified with high sensitivity by means of capillary gas chromatography with electron capture detection. Linoleic acid and sunflower oil were used as model lipid systems. Lipid peroxidation was induced in linoleic acid by the  $\text{Fe}^{2+}$  ion ( $1 \text{ mmol L}^{-1}$ ,  $37^\circ\text{C}$ , 12 h) and in sunflower oil by heating in the presence of  $\text{O}_2$  ( $220^\circ\text{C}$ , 2 h).



U.S. National Library of Medicine

# MMTx API: Process Methods

- MMTx API Java class
  - Use the process method to match your needs:

```
Sentence sentence = mmtxApi.processString ( "obstructive sleep apnea", true );
```

```
public Document processDocument (java.io.File pDocFile)
```

```
public Document processDocument (java.lang.String pDocText)
```

```
public Sentence processSentence(java.lang.String pSentText)
```

```
public Phrase processTerm (java.lang.String pTerm)
```

- Get the output you need from the returned object



U.S. National Library of Medicine



# Future Direction

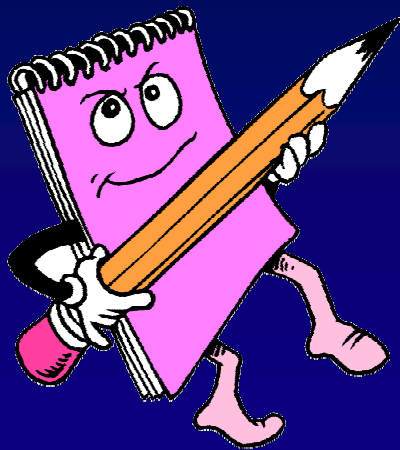


- Bug Fixes
- Bring MetaMap and MMTx closer together
  - Add new MetaMap options to MMTx (*Restrict to Sources, Exclude Sources, Restrict to Semantic Types, Exclude Semantic Types, and Show Preferred Name Sources*).
  - Still have some parsing differences that need to be cleaned up.
- Accept previously POS tagged text
- More detailed output with Phrase and Concept info.



U.S. National Library of Medicine

# Open Session



“Did we get any feedback from management yet?”



U.S. National Library of Medicine

# Our Questions to You

- How are you using MMTx?
- How are you evaluating the results you get?
- Are you already using the Data File Builder or API?
- What would make MMTx work better in your domain?



U.S. National Library of Medicine

# Relevant Links

- MMTx: <http://mmtx.nlm.nih.gov>
- SKR/MetaMap: <http://skr.nlm.nih.gov>
- Indexing Initiative: <http://ii.nlm.nih.gov>



U.S. National Library of Medicine